Paper

Improving interpretability: combined use of LVQ and ARTMAP in decision support

Hoi Fei Kwok, Andrea Giorgi, and Antonino Raffone

Abstract— The learning vector quantization (LVQ) network was used to classify the ECG ST segment into different morphological categories. Due to the lack of data in the ST elevation categories, the classifier was only trained to identify different types of ST depressions (horizontal, upsloping and downsloping). The accuracies were 91%, 85% and 65% respectively for the training, validation and testing data respectively. Despite the low accuracy for the testing data, most of the mis-classifications were downsloping ST depression being classified as horizontal ST depression. We concluded that more data and more training are needed in order to train the LVQ to recognize other morphological types of ST deviation and to improve the accuracy.

Keywords—learning vector quantization, ARTMAP, decision support systems, ischemic heart disease.

1. Background

Cardiovascular diseases have long been recognized to be one of the major causes of morbidity and mortality in developed countries and myocardial infarction is one of the important causes of premature and sudden death in adults. The early detection and diagnosis of myocardial ischaemia allow early intervention and can improve the survival rate. Electrocardiogram (ECG) is an important component of the diagnosis of myocardial ischaemia. In the diagnostic process, the clinicians look at the rhythm of the ECG, the ST-segment deviation, the shape of the segment, the shape of the QRS complexes and the shape of the T wave. Then a decision will be made upon whether the ECG is suggestive of ischaemia supported by the clinical guidelines. Therefore, the diagnosis is usually made in a 2-stage process. Firstly, symptoms, signs and features from the ECG and laboratory investigations are sought. Then, a diagnosis is made using the rules derived from previous research and clinical guidelines.

The computer-assisted diagnosis of ischaemic heart disease has been a subject of research in the past two decades due to an increase in interests in artificial intelligence. However, most systems suffered from a lack of interpretability. Baxt [1] used the presence of a list of medical history, symptoms, signs and ECG features as inputs to an artificial neural network (ANN) but the ECG features were not extracted automatically and no explicit rule could be extracted from the ANN. Other works [3, 4] used different sets of input items but they still required the clinicians to interpret the ECG. Another approach is to use a one-stage approach (for example [2, 7]). The input is the ECG signal and the output is the diagnosis. Neural networks are used to classify the ECG into normal or ischaemic. However, this approach is a black-box approach that does not offer an explanation to the users as to why the system considers the patient having or not having myocardial ischaemia.

In order to improve the interpretability, a rule-based approach is required but in order to use the rule-based approach, one requires information about the inputs to the rules. The inputs to the rules, in the context of the diagnosis of ischaemia from the ECG, include the presence of ST deviation, shape of the ST segment and pathological Q wave, etc. The conventional approach of ST analysis and most standard ECG monitors can give information regarding the magnitude of ST deviation. It is known that ST deviation in itself can be due to other causes [9], therefore, it is important to take the shape of the segment and other information into consideration. However, the recognition of shape cannot be easily done using conventional ECG wave detection method due to the low signal-to-noise ratio. Although one can use a parametric approach by modeling the ST segment by first or second order equations, the segment is short and noisy and the parameters derived can be unreliable. Therefore, a neural network approach is adopted. We used the learning vector quantization (LVQ) network to classify the shape of the ST-segment. The classification can in turn be used as one of the input features to a rule-base system which is developed using the adaptive resonance map (ARTMAP).

2. Overall architecture

Before presenting the use of LVQ in ST analysis, we would like to describe first the overall architecture of the rule-based disease-specific approach of automatic detection of ischaemia from ECG signals. The overall scheme is shown in Fig. 1. The ECG signals are first processed by the ECG processor module. This module filters the ECG and then uses a wavelet-based algorithm to identify the ECG characteristic points (such as positions of R wave, J-point, T wave, etc.) [6]. The module also divides the ECG into segments containing only one beat based on the position of the QRS complexes. These segments can then by used as the inputs to the black-box ST morphology classifier, which is based on the use of the LVQ network.

The ST segment deviation and shape will then be supplied as inputs to the diagnostic rules module. Other features required by the rule-base can be derived from the positions of ECG characteristic points. These features include heart rate, T-wave inversion, tall R wave, widened QRS complex and pathological Q wave, etc.

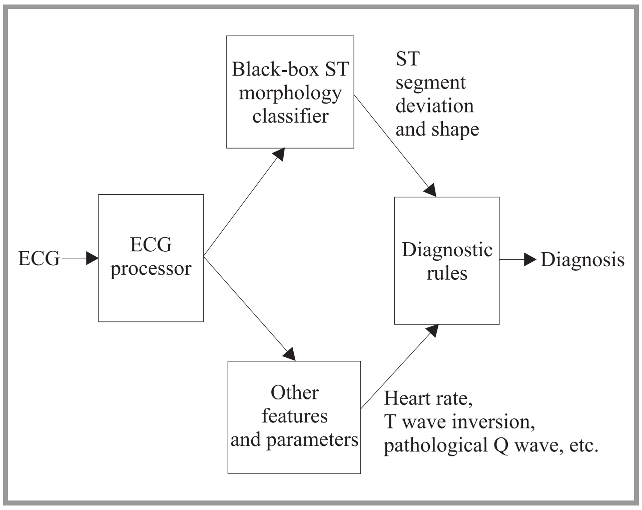


Fig. 1. The overall scheme for the ischaemic classification from ECG. The black-box ST morphology classifier is based on the LVQ network and the "diagnostic rule" module is based on ARTMAP.

The diagnostic rules module consists of a set of "if-then" rules. These rules were derived using the ARTMAP. A graphic user interface (GUI) was implemented and used for presenting the ECG to the cardiologists and recording their classification. ECG segments were presented to the cardiologists and they were asked to indicate on the screen the presence or absence of the input features and the diagnosis (i.e., normal, probably ischaemia, ischaemia, abnormal but not ischaemia, etc.). The data were then used for the ARTMAP training. After training, the rules were derived from the top-down weights of the ART_a module and were reviewed by the cardiologists. Unacceptable rules were excluded. In this paper, we will not go into the details of the ARTMAP as the primary focus is the use of LVQ.

3. Learning vector quantization

The LVQ network [5] is a neural network related to the vector quantization and can be used for pattern classification. Here we will only present the LVQ1 which was used in our system. The LVQ1 consists of an input layer and an output layer. The dimension of the input layer is the same as that of the input vector. Each node in the output layer represents one output class. The activity of each output node is the Euclidean distance between the input vector and the input weight vector. The output layer is a competitive layer and thus, the output from the node with the highest activity will be 1 and the outputs from the rest will be zeros.
During the training, the weight vector is adjusted according to the output class and the target class. The target class is the desirable target. Let $\mathbf{w_c}$ be the input weight vector of the output class node, $\mathbf{x}$ be the target class. The weights are adjusted according to the following equations:

$$\mathbf{w_c}(n+1) = \mathbf{w_c}(n) + \alpha(n)\left[\mathbf{x}(n) - \mathbf{w_c}(n)\right] \quad \text{if } \mathbf{x} \text{ and } \mathbf{w_c} \text{ belong to the same class,} \tag{1}$$

$$\mathbf{w_c}(n+1) = \mathbf{w_c}(n) - \alpha(n)\left[\mathbf{x}(n) - \mathbf{w_c}(n)\right] \quad \text{if } \mathbf{x} \text{ and } \mathbf{w_c} \text{ belong to different classes,} \tag{2}$$

where $\alpha(n)$ is the learning rate at the nth epoch. The input weights of other nodes remain unchanged. Using this algorithm, the input weight vector will get closer to the input vector as time progresses.
The network is modified in the MATLAB neural network toolbox. An additional layer of neurons is added to the output layer so that the output layer becomes the hidden layer. The number of nodes in this new hidden layer can be greater than the number of output classes (named subclasses). The dimension of the new output layer is the same as the number of output class. Therefore, a few subclasses can be associated with the same output class using a binary connection weight. This is the network structure that was adopted in this paper.

4. Methodology

4.1. Training and testing data collection

The ECG signals from the European ST-T database were used. Only those from V4 were used in the study as not enough data can be retrieved for the training purpose from other leads. The QRS complexes in the ECG signals were located and used as reference points for the truncation of the signals into individual beats. Each individual beat was presented to a clinician for the classification of ST morphological type. For the development and training of the LVQ network, each input-target data set consisted of the time-series voltage values from one ECG beat and the corresponding ST-morphology assigned by the clinician. The data were divided into training data set, cross-validation data set and testing data set.

4.2. Data preprocessing and network training

Each ECG beat was re-sampled to give 128 data points. The signals were then normalized to zero mean and unity standard deviation. The principal component analysis (PCA) was then applied to the data and the coefficients each principal component were used as the input to the LVQ network. During the PCA, only components accounted for more than 1% of the total variance were included. In this way, the input dimension was reduced to 15 from 128. LVQs with 16 to 30 hidden nodes were used and the LVQ which showed the best performance was chosen. The training was done in series of 100 epochs until a maximum of

1000 epochs. After every 100 epochs, the prediction error in the cross-validation data set was calculated. The training stopped when an increase in the prediction error was detected.

5. Results

The 486 beats were analyzed by the clinician; 7 beats were unclassified (due to large signal noise) and only 37 out of 486 beats showed ST elevation. Therefore, a decision was made not to include ST elevation as one of the output categories. The LVQ so developed could therefore only classify the ECG beat into "normal", "horizontal ST depression", "upsloping ST depression" and "downsloping ST depression". The LVQ with 27 hidden nodes was found to give the best performance. It gave the correct classification in 91% of the training data, 87% of the cross-validation data and 65% of the testing data (Table 1). Most of the mis-classifications in the testing data were classified as "downsloping ST depression" by the clinician and the network classifications were "horizontal ST depression".

Table 1
The performance of the LVQ network
with 27 hidden nodes

Target	LVQ output											
	Training data				Validation data				Testing data			
	1	2	3	4	1	2	3	4	1	2	3	4
1	31	2	0	1	25	3	2	1	49	6	0	0
2	9	37	1	6	9	36	3	6	5	17	1	2
3	0	4	40	1	3	7	36	4	3	0	0	0
4	0	8	0	30	0	7	0	28	1	15	2	0
Target:	1 – normal, 2 – horizontal ST depression, 3 – upsloping ST depression, 4 – downsloping ST depression.											

6. Discussions and conclusions

The accuracy of the trained network on the training data and the validation was satisfactorily however, the accuracy on the testing data was poor. However, when one looks into the details of the make-up of the testing data, one can see that the data were not evenly distributed. Only three cases were classified by the clinician as upsloping ST depression. The mis-classification of downsloping ST depression into horizontal ST depression may indicate that the class boundary should be shifted more towards the downsloping ST depression. The performance may be improved by further training. Another solution is to use the LVQ2. In LVQ2, not only the weight vector closest to the input will be modified but the one second closest to the input may also be modified under defined circumstances.

As mentioned earlier, more data will be needed in order to train the neural network to recognize different types of ST elevation and to recognize the ST morphology in more leads. The LVQ network was chosen in this paper for this application because each ST-segment can only belong to one morphological class and the LVQ network can be trained to discover the class boundaries and the competitive layer ensures that only one class is chosen as the output. Despite this, the use of other types of neural network can be investigated in the future.

References

[1] W. G. Baxt, "Use of artificial neural network for the diagnosis of myocardial infarction", *Ann. Int. Med.*, vol. 115, pp. 843–848, 1991.

[2] B. Heden, M. Ohlsson, R. Rittner, O. Pahlm, W. K. Haisty Jr, C. Peterson, and L. Edenbrandt, "Agreement between artificial neural networks and experienced electrocardiographer on electrocardiographic diagnosis of healed myocardial infarction", *J. Amer. Coll. Cardiol.*, vol. 28, pp. 1012–1016, 1996.

[3] J. S. Jorgensen, J. B. Pedersen, and S. M. Pedersen, "Use of neural networks to diagnose acute myocardial infarction. I. Methodology", *Clin. Chem.*, vol. 42, pp. 604–612, 1996.

[4] R. L. Kennedy, R. F. Harrison, A. M. Burton, H. S. Fraser, W. G. Hamer, D. MacArthur, R. McAllum, and D. J. Steedman, "An artificial neural network system for diagnosis of acute myocardial infarction (AMI) in the accident and emergency department: evaluation and comparison with serum myoglobin measurements", *Comput. Meth. Progr. Biomed.*, vol. 52, pp. 93–103, 1997.

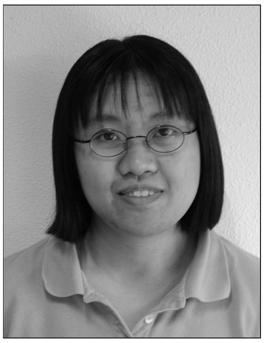
[5] T. Kohonen, "Learning vector quantization", in *Self-Organising Maps*, Springer Series in Information Sciences. Berlin, Heidelberg, New York: Springer-Verlag, 1995, 3rd ed.

[6] H. F. Kwok, A. Giorgi, R. Fenici, and A. Raffone, "Identification of electrocardiogram characteristic points: wavelet transform vs derivative-based method", *J. Amer. Coll. Cardiol.*, vol. 43, no. 5 (Suppl. A), p. 400A, 2004.

[7] N. Maglaveras, T. Stamkopoulos, and M. G. Strintzis, "An adaptive backpropagation neural network for real-time ischemic episodes detection: development and performance analysis using the European ST-T database", *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 805–813, 1998.

[8] B. P. Simon and C. Eswaran, "An ECG classifier designed using modified decision based neural networks", *Comput. Biomed. Res.*, vol. 30, pp. 257–272, 1997.

[9] K. Wang, R. W. Asinger, and H. J. L. Marriot, "ST-segment elevation in conditions other than myocardial infarction", *New Eng. J. Med.*, vol. 349, pp. 2128–2135, 2003.



Hoi Fei Kwok received the M.B.Ch.B. degree from the Chinese University of Hong Kong, the M.Sc. degree in medical physics and clinical engineering from the Sheffield University, UK, and the Ph.D. degree in automatic control and systems engineering from the Sheffield University, in 1991, 1998 and 2003, respectively. She also received the FRCA from the Royal College of Anaesthetists, London, UK in 1999. She worked as a research associate in the Department of Automatic Control

and Systems Engineering, University of Sheffield from 1999 to 2003. As part of the EU funded HEARTS project, she joined the Department of Psychology in Sunderland University in 2003. In 2004, she worked as a visiting research scientist in the RIKEN Brain Science Institute, Japan. Currently she is a research fellow in the Behavioural Neuroscience Centre, School of Psychology, University of Birmingham, UK. Doctor Kwok research interests include neural networks, neuronal modelling, fuzzy logic, human motor control, wavelets and decision support systems in biomedicine.

e-mail: h.f.kwok@bham.ac.uk
Sensory and Motor Neuroscience (SyMoN)
Behavioural Brain Science Centre
Hills Building
The University of Birmingham
Edgbaston
Birmingham B15 2TT
United Kingdom

Andrea Giorgi was born in Rome, Italy, in 1972. He obtained his Medical School degree in 1996 at the "La Sapienza" State University of Rome, and specialization degree in internal medicine in 2002 at the Catholic University of Rome Medical School. He qualified for medical profession in 1997. Since 1998 he collaborates with the Clinical Physiology and Biomagnetism Research Center of the Catholic University of Rome and the Biomag Center of the Helsinki University of Technology, and will attain a Ph.D. in "Advanced technologies in clinical physiology and cardiac multimodal imaging", in 2005. He collaborates with the Department of Cardiovascular Medicine and the Institute of Clinical Chemistry of the Catholic University of Rome for the study of cardiac markers in congestive heart failure. Since 2003, he also collaborates with the Biomath Lab (I.A.S.I.-C.N.R), for the development of the IST project "HEARTS" (health early alarm recognition and telemonitoring system). At present he is performing his clinical practise at the Department of Internal Medicine of the San Raffaele Hospital – Rome – Tosinvest Sanità. Doctor Giorgi participated with papers of his own in several national and international congresses and published, as author and co-author, on national and international journals. His interests include magnetocardiography, neural networks in cardiology, and markers in cardiology.

e-mail: akgiorgi@tiscali.it
Department of Internal Medicine, Sensory
and Motor Neuroscience
San Raffaele Hospital (IRCCS)
Via della Pisana n. 237 cap 00100
Rome, Italy

Antonino Raffone received M.Sc. degree in psychology at the University "La Sapienza" of Rome in 1995 and completed specialisation in "Cognitive Psychology and Neural Networks" at the same university in 1996. He also completed a Ph.D. in "Cognitive Psychology" in 2002, at the same university, and received the European diploma in cognitive and brain sciences, from 2000 to 2002. He was a lecturer at the University of Sunderland (UK) from 2001 to 2005, and has recently started a professorship in psychology (teaching "Neurocognitive Modelling") at the University "La Sapienza" of Rome, with a "Re-entry of the Brain" national award. Doctor Raffone is visiting researcher at the Laboratory for Perceptual Dynamics of the RIKEN Brain Science Institute in Japan, and member of the Interuniversity Centre for Research on Cognitive Processing in Rome. His research interests include computational cognitive neuroscience (with special reference to prefrontal cortex and hippocampus), synchronization dynamics and psychophysics of interactions between visual attention and working memory.

e-mail: antonino.raffone@sunderland.ac.uk
Theoretical and Applied Simulation Laboratory
Department of Psychology
Sunderland University
United Kingdom